REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)

2. REPORT DATE
30 December 1996

3. REPORT TYPE AND DATES COVERED
Technical (14 Nov 1994 - 31 Dec 1995)

4. TITLE AND SUBTITLE
Use of Wavelets to Extract Higher Order Statistics for Separation of Targets and Clutter

5. FUNDING NUMBERS
DAAH04-95-C-0003

6. AUTHOR(S)
James Meany and Steven Wright

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
McDonnell Douglas Aerospace
P.O. Box 516, Mailcode S270-1115
St. Louis, MO 63166

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARO 33584.5-MA

11. SUPPLEMENTARY NOTES
This work was performed on a contract under the ARPA University-Industry ATR Initiative.

12a.. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE
N/A

13. ABSTRACT
Summarizes work performed in 1995 by McDonnell Douglas Aerospace on a contract under the ARPA University-Industry ATR Initiative. The focus is on an investigation into the use of wavelets to extract higher order statistics for separation of targets from clutter in sensor imagery.

14. SUBJECT TERMS
Multiresolution Analysis, Wavelets, Automatic Target Recognition

15. NUMBER OF PAGES
13

16. PRICE CODE
N/A

18. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	18. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

USE OF WAVELETS TO EXTRACT HIGHER ORDER STATISTICS FOR SEPARATION OF TARGETS AND CLUTTER

TYPE OF REPORT: TECHNICAL

JAMES MEANY AND STEVEN WRIGHT

30 DECEMBER 1996

U. S. ARMY RESEARCH OFFICE

CONTRACT / GRANT NUMBER: DAAH04-95-C-0003

MCDONNELL DOUGLAS AEROSPACE
P.O. BOX 516, MAILCODE S270-1115
ST. LOUIS, MO 63166

APPROVED FOR PUBLIC RELEASE;

DISTRIBUTION UNLIMITED

TABLE OF CONTENTS

LIST OF FIGURES

1. Background

This report describes an investigation into the use of wavelet coefficients to extract higher order statistics from sensor data, in an effort to separate mobile targets from clutter. This investigation was conducted in the context of McDonnell Douglas' General Pattern Match (GPM) ATR algorithm, which provides a decision theoretic approach to target recognition and has been successfully used on several programs and with a variety of sensors, including synthetic aperture radar, laser radar, forward-looking infrared (FLIR), and millimeter wave.

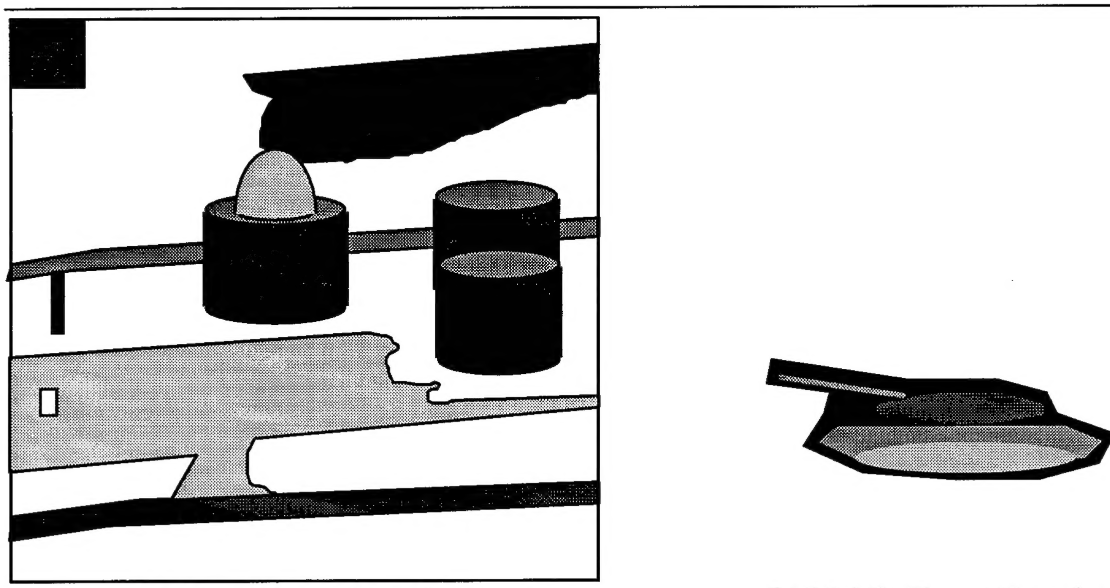
1.1 Probability Density Functions for Sensor Data

The GPM approach results from a rigorous application of optimal decision theory [1], but does not rely on knowledge of the probability density function for the sensor data nor on assumptions of Gaussian statistics. This is a crucial step because the random processes governing image data have proven to be highly non-Gaussian and are often unknown or unpredictable (as in FLIR imagery, for example). When the probability density function for a sensor is unknown, the density function is estimated from the sensed image data itself.

1.2 Target References

Target shapes are represented using partitioned image templates to represent the appearance or shape of target or scene objects.. A target template defines the spatial shape of significant pixel population groups, by assigning arbitrary region codes to subelements of a target or scene object. For example, a tank template may have region codes for turret, treads, and body; or a building may have region codes for walls and roof. The region codes designate which pixels should be grouped into common distributions for the purpose of computing likelihood scores. The use of target shape information (in the form of partitioned templates) for evaluation of target hypotheses is based on an implicit assumption that the pixel intensities in sensed imagery derive from physical scene attributes such as shape, orientation, temperature, illumination, and material composition. The pixel intensities for scene elements sharing similar physical attributes are typically governed by similar statistics.

Because target references are represented using partitioned image templates, target reference information can readily be derived from either 3-D physical target models or from reconnaissance imagery (in a so-called "photo matching" mode). The choice between photo matching and 3-D model matching approaches provides a unified yet versatile decision theory approach to a variety of target recognition applications. For fixed land targets, the target template consists of a partition of the target and its surrounding context into regions, rendered to the 2-D perspective view of the target from the estimated approximate position of the weapon system. Such a template may be derived either by rendering a 3-D site model of the target to the 2-D perspective view or by "warping" a reconnaissance image augmented with selected 3-D information. For mobile or relocatable targets, the target template is obtained by rendering a 3-D facet model of the target to the 2-D perspective view at the hypothesized range/scale and orientation of the target. This rendered mobile target template is combined at a hypothesized position with any background regions estimated from the sensed image data to represent the hypothesis in the form of a complete image partition. Sample target templates for a fixed land target and a mobile target are shown in Figure 1.



(a) Fixed Target Template (b) Mobile Target Template

Figure 1. Sample Target Templates

1.3 Likelihood Function Evaluation

For the evaluation of a target hypothesis, a hypothesized image partition (representing target shape an d position information for the target hypothesis) is applied to the sensed image. A likelihood function calculation generates a match score measuring how well the regions of hypothesized image partition match the sensed image data. Figure 2 provides a notional depiction of this process. If phenomenological modeling of the sensor, target, and background is available, the modeled probability distributions of pixel intensities are included in the likelihood function calculation. The likelihood functions are actually represented as log-likelihoods which simplifies computation by changing multiplication operations for the evaluation of joint probabilities to summations. Sensor fusion applications are easily supported by simply summing the match scores obtained from the two sensors.

The likelihood function calculation includes evaluation of a counter-hypothesis based on the absence of a target. The log-likelihood for the counter-hypothesis is subtracted from the log-likelihood for the hypothesis. Inclusion of the counter-hypothesis calculation eliminates false alarms on certain background regions which would otherwise yield a good match to the target model. The counter-hypothesis can include a clutter model estimated from the image data.

1.4 Search Strategy

The evaluation of target hypotheses proceeds in accordance with a coarse-to-fine layered search strategy, supporting target detection, recognition, and identification, with these three functions differentiated only by the resolution of the search in the search space dimensions of position, orientation, range, and target type. For mobile targets, target detection is performed at coarse resolution using a generic target detection template not specific to any target type. The generic target detection template is generated to match the peak envelope of all target types at all aspects at a specified range. An example of a target detection template is shown in Figure 3. Such a strategy

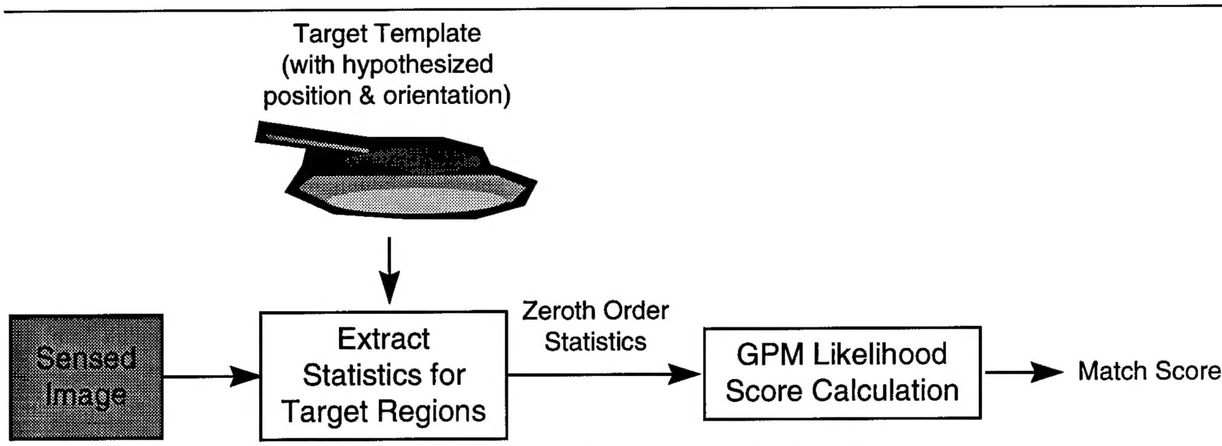


Figure 2. Likelihood Function Evaluation Process

is aimed at separating candidate targets from clutter, with the best matches nominated to the next layer of the search for further processing.

Computational throughput or time limitations are enforced by limiting the number of candidates nominated through the search process. At successive layers, the search proceeds at increasing resolution. The final stage of the search provides target identification with the position, orientation, range, and classification of any targets fully resolved.

2. The Problem: Limitations of Zeroth Order Statistics

In the approach described above, target decisions are based on likelihood scores computed from zeroth order pixel statistics, represented in the form of a histogram extracted for each region of a hypothesized target template which has been applied to the sensed image data. Because the zeroth order analysis is invariant to spatial redistributions of pixels within a region, this approach is insensitive to the unmodeled internal structure of a region and also to texture within a region. Unmodeled internal structure or texture within a region can reduce the likelihood score for the region and increase the probability of false alarms and false matches on clutter and non-target objects.

The reliance on zeroth order pixel statistics corresponds to an assumption that the pixel intensities are statistically independent from pixel to pixel, which causes any information associated with dependencies between pixels to be discarded from the target decision. In practice, neighboring pixels which derive from the same physical surface typically do exhibit dependencies, and this information can be critical to target recognition decisions. For example, we may have a case in which the raw distributions of pixels from a target and background are identical, but the target is still distinguishable due to a distinct change in the texture pattern. The direct statistical approach to modeling dependencies between pixels would be to represent the observed data as a joint density function of a neighborhood of several pixels. Differences in texture patterns and internal structure would then result in distinctly different joint pixel distributions. However, this approach is not practical for two reasons:

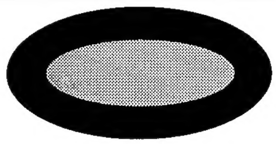


Figure 3. Target Detection Template

- The joint distribution of an n pixel neighborhood of 8 bit pixels is an n dimensional histogram with 2^{8n} distinct bins. The joint distribution for a pixel pair would have 2^{16} bins, while the joint distribution for just a 5 pixel neighborhood would have 2^{40} bins. Handling such large distributions is computationally prohibitive using current and even foreseeable technology.

- A more serious theoretical problem is that the large joint distributions would be too sparsely populated to produce meaningful statistics.

One approach for overcoming the limitations of zeroth order statistics is to utilize target templates with high levels of detail. In the case of fixed land targets, the use of 3-D warped reconnaissance imagery for target templates provides sufficient detail of the target structure to effectively support the identification of the target aimpoint in clutter. With this "photo matching" approach, clutter in the vicinity of the target is included in the target reference and actually serves as an additional source of information for matching the target. In the case of mobile targets, a detailed 3-D facet model of the target rendered to the hypothesized 2-D perspective often provides sufficient detail to separate the target from clutter.

A remaining area of difficulty lies in the detection of mobile targets in clutter using a generic target detection template such as is shown in Figure 3. While the use of detailed target templates could provide improve the separation of targets from clutter, throughput limitations preclude the use of detailed templates at early (detection) stages of the search process. Unfortunately, the lack of details in the generic target detection template makes the detection process vulnerable to an increased probability of false alarm and a reduced probability of detection.

3. Extraction of Higher Order Statistics Using Wavelets

An alternative approach for handling the higher order statistics in the sensed image data is to identify a transformation of the image pixel data that reduces the joint random distribution to a distribution of a single new random variable. Differing texture patterns in the original image data would be detected by a change in the intensity distribution of transformed pixel values. With a similar motivation, we have in the past successfully used the Sobel edge transform as a method for capturing first-order joint pixel distribution statistics from sensor data. This use of the Sobel transform information increased the probability of acquisition by 35% with co-boresighted laser radar and FLIR sensor data during the Cruise Missile Advanced Guidance (CMAG) program (detailed results are classified).

Because the wavelet transform is spatially and spectrally localized, it is well-suited to isolate texture patterns and other structurally significant higher order statistical dependencies. The power of combining wavelet processing with the our decision theoretic approach is that we do not require the wavelet coefficients to be Gaussian-distributed random variables. The evaluation of the likelihood functions can utilize any resulting coefficient distribution so long as the wavelet

transform produces coefficients that have a different distribution of values for target and background data.

When dealing with likelihood function theory, there is a natural method for combining information from multiple sources: summation of the log-likelihood functions computed independently for each of the information sources. If the information sources are statistically independent, this combination is optimal. Even when the information sources are not strictly independent, summation of the log-likelihood results works well because the correct non-Gaussian marginal distributions are being used for the decomposed problem. This is unlike a typical application of log-likelihood's in which a Gaussian form is assumed for the marginal distributions in addition to the assumption of independence.

Figure 4 illustrates the use of wavelet transforms to extract higher order image statistics. A wavelet transform is applied to the sensed image and then the process of using the target template to partition an image and the computation of log-likelihood functions can be applied to the original image as well as one or more selected subbands of the wavelet transform. The individual log-likelihood scores may then be summed to create a single log-likelihood. The result is a single numerical value upon which target decisions can be made.

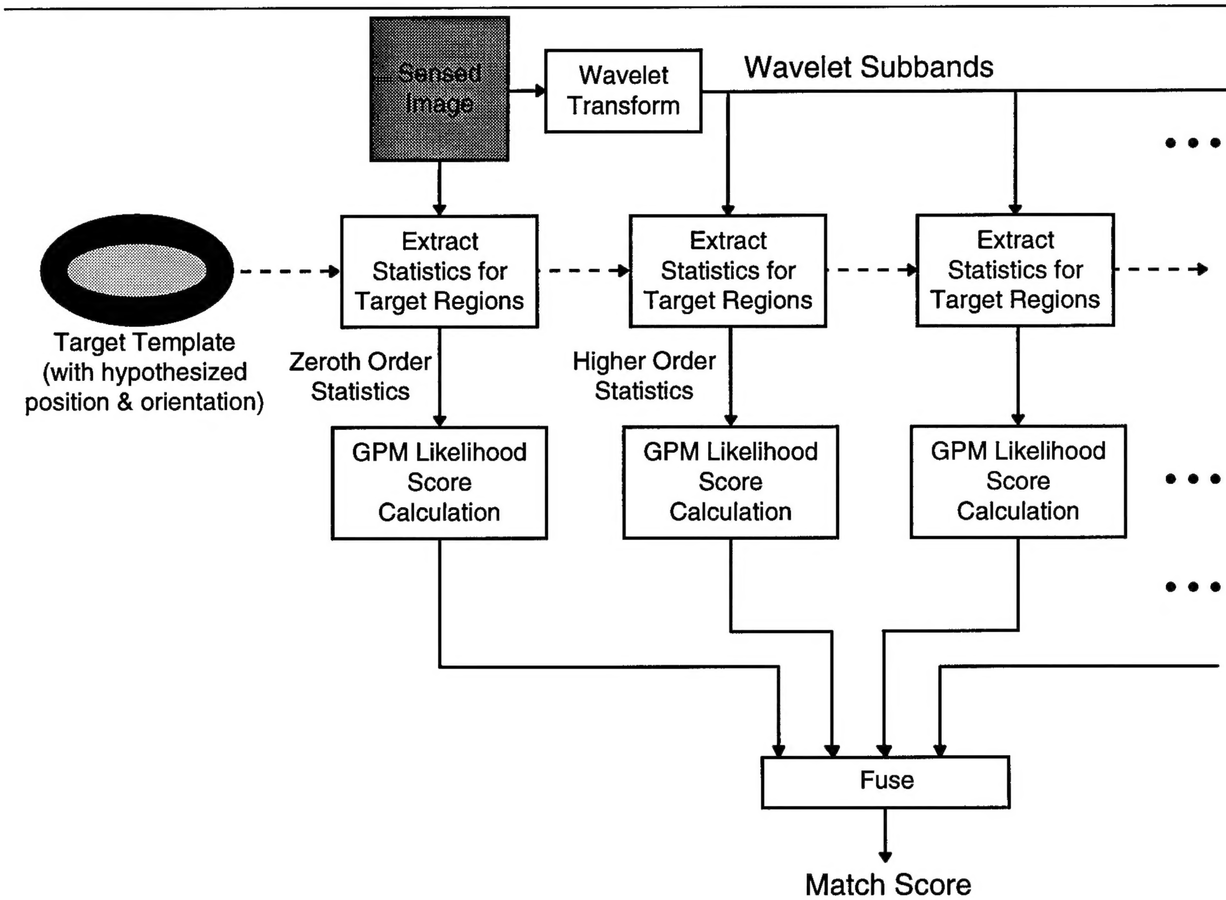


Figure 4. Use of Wavelet Transform to Extract Higher Order Statistics

4. Target Screening Investigation

To investigate the use of wavelets for extraction of higher order image statistics, we implemented the approach of Figure 4 in the context of the GPM Engine, a real-time hardware implementation of McDonnell Douglas' GPM algorithm. The wavelet transform was implemented using biorthogonal wavelets with the scaling filter and wavelet filter having lengths of 5 taps and 3 taps respectively.[2] The original image and the wavelet subbands at the finest resolution level were used in the computation of the likelihood scores. The finest band was chosen because the filters are more spatially localized than at coarser resolutions.

This configuration was applied to the problem of screening (detection) of mobile targets, using FLIR imagery from a Loral 3-5 micron platinum silicide focal plane array of 512x474 elements with a field of view of 3 degrees by 2.25 degrees. The target set consisted of a tank, a truck, and an armored personnel carrier. The target approaches included broadside and end-on approaches in high and low clutter. Figure 5 shows an example image of the targets at medium range. The target screening was performed using a generic elliptical target detection template generated to match the peak envelope of all target types at all aspects at a hypothesized range. The detection search strategy was configured to designate the best matches which could then be further processed at finer resolution for the search dimensions and with more detailed target templates.

Figures 6, 7, and 8 plot the probability of detection versus range to target under various clutter and target approach directions, and with detection list lengths of 10, 5, and 3, respectively. These results show conclusively that the use of wavelets does not improve the target screening performance, and usually actually degrades performance.

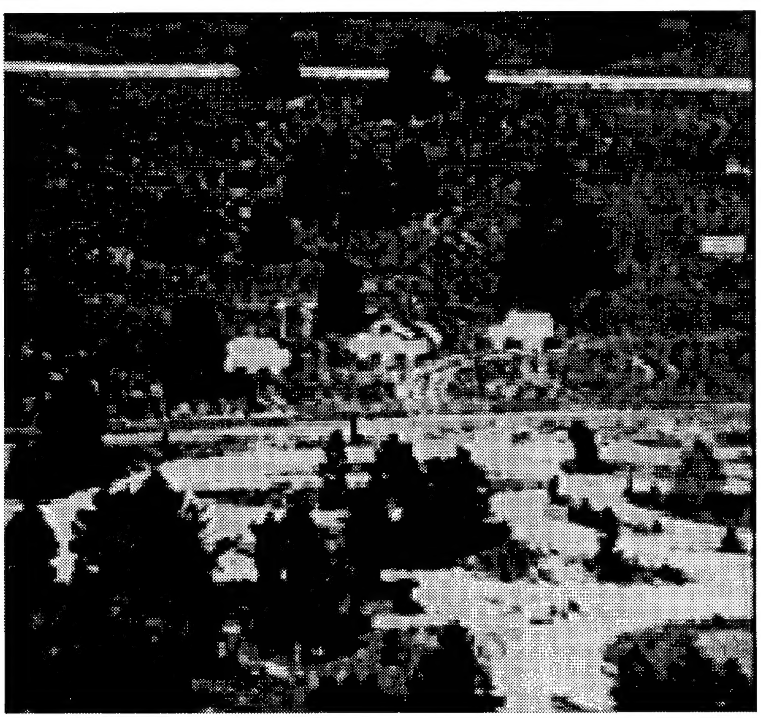


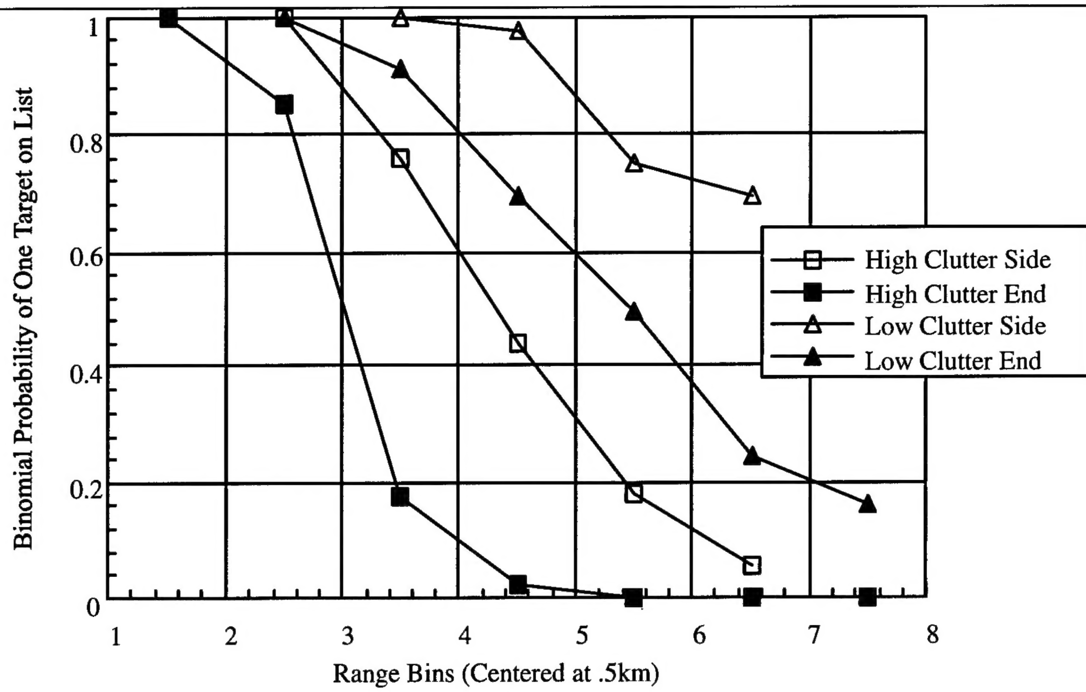
Figure 5. FLIR Image of Armored Personnel Carrier, Tank, and Truck

A detailed examination of the results revealed that the interaction of the wavelet filters with the boundaries of objects in the scene degraded the likelihood scores. Because the wavelet filters are highpass, they respond strongly to image edges. As a result, the wavelet coefficients exhibit different statistics along an edge than on either side of the edge. The mixture of the statistics for wavelet coefficients within regions and along boundaries causes the likelihood scores to be degraded. These problems are not alleviated by the use of alternative wavelet filters since all wavelet filters are highpass and respond to edges. Also, the use of coarser subbands exacerbates the problem because the coarser wavelet filters have poorer spatial localization.

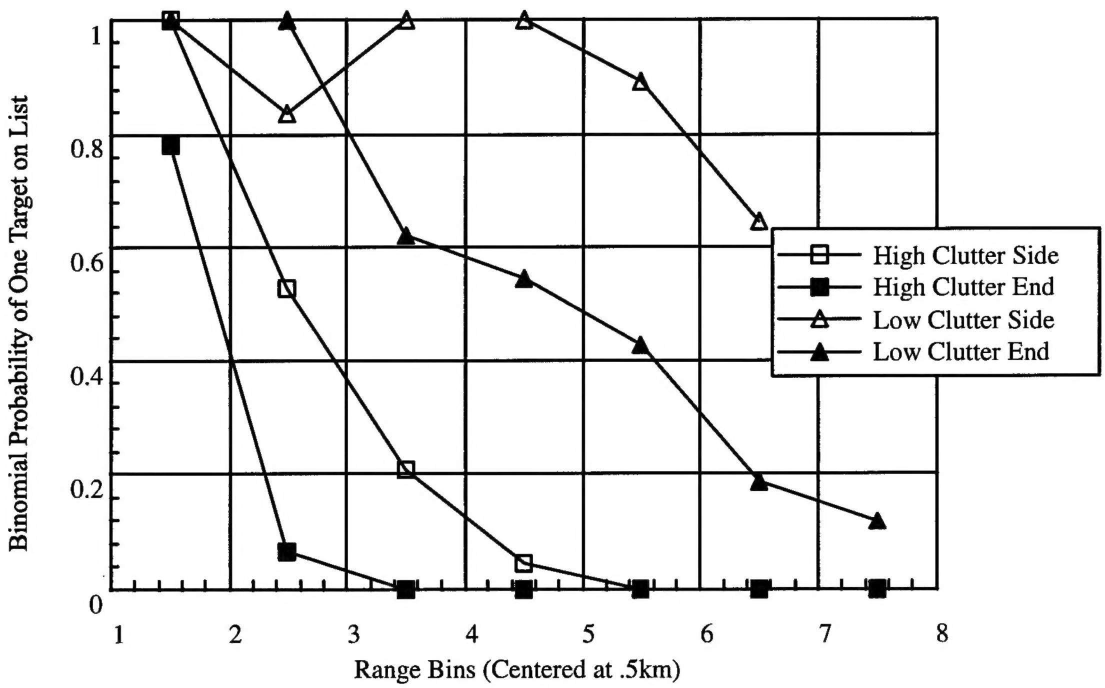
A possible remedy to this problem would be to create a separate region or regions along the boundaries of the target templates. The boundary region(s) would then isolate the statistics of the boundaries from the statistics of the interiors of adjacent regions. We have successfully employed this measure in the past when using the Sobel transform in conjunction with model-based target recognition. However, this approach is of no use when applying a generic target detection template because the generic shape of the template lacks sufficient detail to accurately match the true boundaries of targets in the scene. Unless the boundaries of the template accurately match the detailed shape of actual scene objects, the statistics of coefficients at the boundaries are not effectively separated from the statistics of coefficients within adjacent regions. On the other hand, as indicated earlier, when a detailed target template is employed, there is little or no need for the extraction of higher order statistics from the image because the use of the detailed template typically allows a target to be effectively separated from clutter. Unfortunately, the use of detailed target templates at the early detection stages of a layered search strategy is computationally prohibitive.

Based on these results, we have determined that the use of wavelets to extract higher order image statistics in conjunction with generic target detection templates is not effective, due to the boundary effects discussed above. We are continuing to investigate other uses of wavelets in automatic target recognition. Specifically, we have begun an investigation into the use of wavelets to extract local feature points ("tie points") to be used for matching 2-D and 3-D structure in multiple images. The approach is to detect dominant local discontinuities (i.e., point features, corners, curvature discontinuities) as a function of wavelet coefficients. Potential applications include registration of 2-D imagery, estimation of 3-D geometric structure from image sequences for matching to target references, and extraction of 3-D structure from reference imagery for use in reference preparation and mission planning. Results of this investigation will be included in subsequent technical reports and manuscripts.

In another related effort under this contract, Summus Ltd. is investigating the use of wavelets to obtain a local oscillation measure related to fractal dimension. This will be used to differentiate target and clutter objects based on their texture properties. The results of this effort are documented in a separate report by Summus.

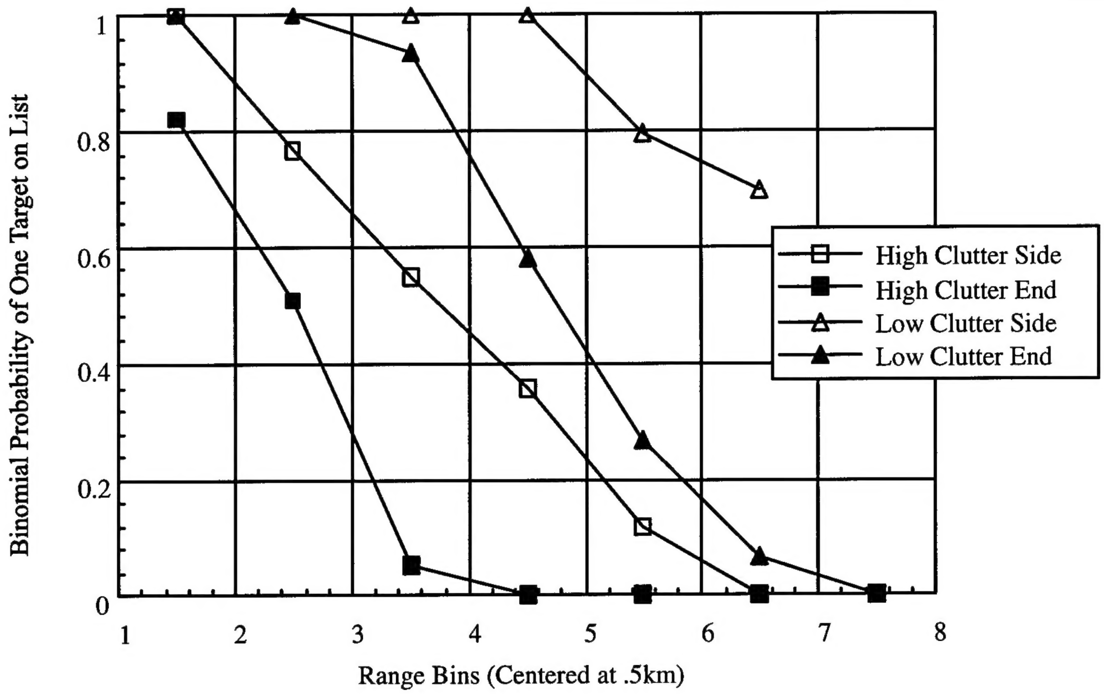


(a) Results for GPM Screener (without Wavelets)

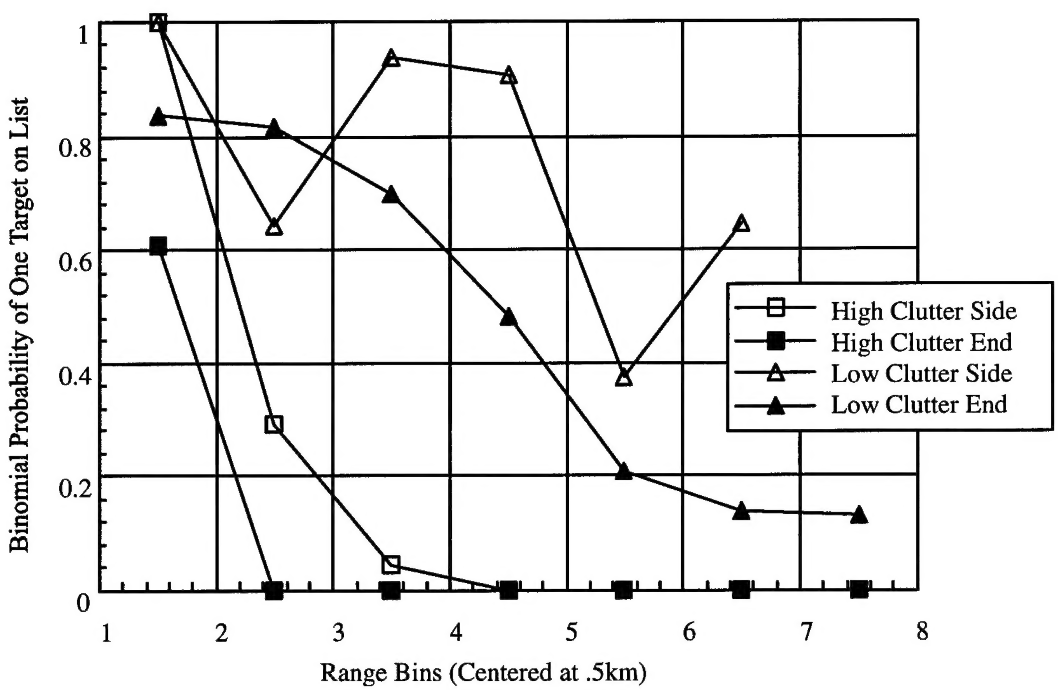


(b) Results for GPM Screener (with Wavelets)

Figure 6. Target Screening Results Using List Length 10

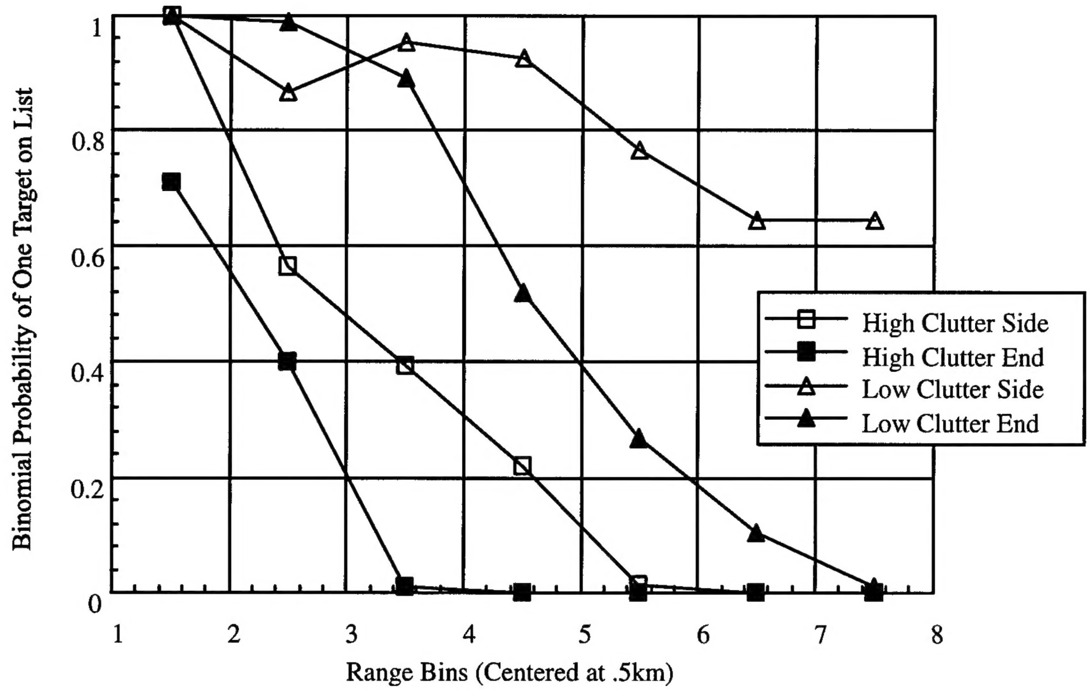


(a) Results for GPM Screener (without Wavelets)

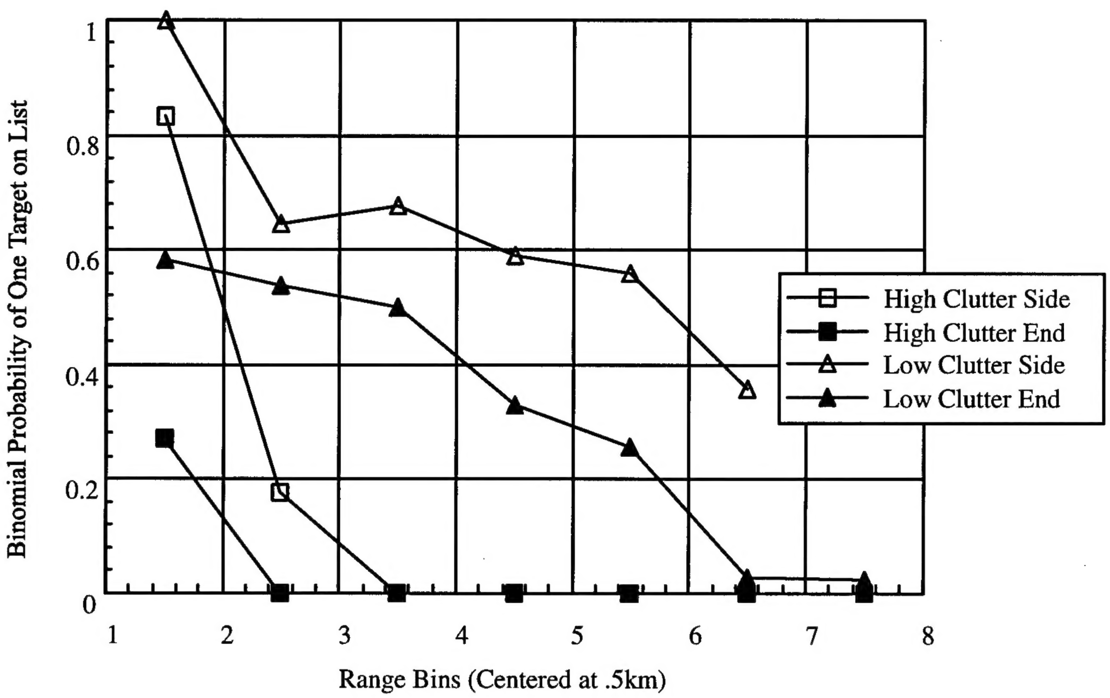


(b) Results for GPM Screener (with Wavelets)

Figure 7. Target Screening Results Using List Length 5



(a) Results for GPM Screener (without Wavelets)



(b) Results for GPM Screener (with Wavelets)

Figure 8. Target Screening Results Using List Length 3

References

1. Harry L. Van Trees, *Detection, Estimation, and Modulation Theory*, John Wiley & Sons, 1968.

2. A. Cohen, "Biorthogonal Wavelets", *Wavelets – A Tutorial in Theory and Applications*, C.K. Chui (ed.), Academic Press, 1992.